

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
11 April 2002 (11.04.2002)

PCT

(10) International Publication Number  
**WO 02/29610 A2**

(51) International Patent Classification<sup>7</sup>: **G06F 17/10**

Hobart Street, Pittsburgh, PA 15217 (US). **PTAK, Andrew**; 1908 Shaw Avenue, Pittsburgh, PA 15217 (US).

(21) International Application Number: PCT/US01/31164

(22) International Filing Date: 4 October 2001 (04.10.2001)

(74) **Agent: JAMES, Richard, W.**; Morgan, Lewis & Bockius LLP, 1800 M Street, N.W., Washington, DC 20036-5869 (US).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
09/679,694 5 October 2000 (05.10.2000) US

(81) **Designated States (national)**: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(71) **Applicant: DIGITALMC CORPORATION** [US/US]; 5543 Wilkins Avenue, Pittsburgh, PA 15217 (US).

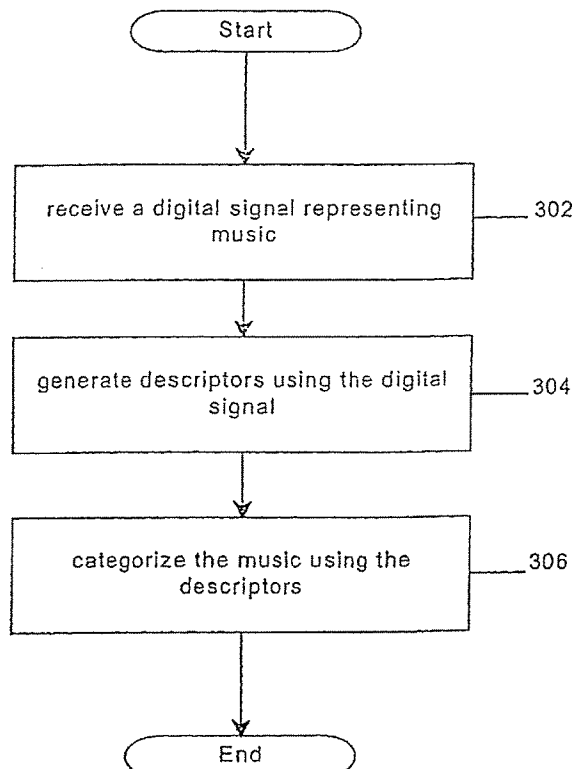
(84) **Designated States (regional)**: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European

(72) **Inventors: BANKS, Annette, P.**; 5543 Wilkins Avenue, Pittsburgh, PA 15217 (US). **NICHOL, Robert, C.**; 5419

[Continued on next page]

(54) Title: METHOD AND SYSTEM TO CLASSIFY MUSIC

300



(57) **Abstract:** A method and apparatus for categorizing music is described. A digital signal representing music is received. Descriptors are generated using said digital signal. The music is categorized using said descriptors.



WO 02/29610 A2



patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declarations under Rule 4.17:**

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for all designations*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*

**Published:**

- *without international search report and to be republished upon receipt of that report*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

## METHOD AND SYSTEM TO CLASSIFY MUSIC

FIELD OF THE INVENTION

The present invention relates to classifying recorded music into categories. In particular, the system and method of the present invention provide a unique method of  
5 classifying music using digital signal analysis.

BACKGROUND OF THE INVENTION

Sales of digital music over the Internet are increasing rapidly. By 2007, sales of music over the Internet are projected to approach \$4 billion a year. This increase in sales is being driven by technology. As computers become larger and faster, more  
10 data can be stored and quickly analyzed. Rapid growth in high speed Internet access in homes through ADSL, DSL, wireless and cable modems is driving the rapid growth in digital downloads of music.

In addition to improvements in the hardware, there have been significant developments in new software applications. One example is the development of the  
15 MP3 music data format. This standard is a method of storing musical data that reduces the storage size of the information to a tenth of its original size, thus facilitating the rapid download over the Internet. New hardware development, such as portable Internet radios and MP3 walkmans, and new software initiatives, such as MPEG-4 and SDMI, are currently underway. These will also contribute to the growth  
20 of the downloaded music industry.

Access to music over the Internet allows people to have access to all types of music. The Internet's innate qualities of searchability, convenience and cost savings will make it the predominant medium for music delivery in the future.

This type of widespread access to every type of music imaginable changes the  
25 sales strategy of the music industry. Previously, the music industry decided what music people desired to listen to through strategic CD advertising and radio station playlists. Before the advent of the Internet, musical artists without recording contracts were generally unable to sell their music on a widespread basis. However, access to

music over the Internet means that people can download and purchase many types of music that may not have been available in traditional formats. The Internet has been, and will continue to be, an incredible opportunity for unestablished artists to sell their music.

5           As the popularity of downloading music increases, there are problems for both companies in the music industry and for buyers of downloaded music. For the music industry, there are concerns relating to the ability to reach customers as retail store sales decline. For consumers, the concern is how to find the music that they like, particularly as many new artists make their offerings available for free and established  
10 artists increasingly attempt to sell their music directly to the consumer.

Presently, in order to assist a customer in finding the type of music he wants, music is classified into a number of different categories. Typically, most methods of classifying music involve subjectively categorizing music into one of a number of genres, such as blues, rock, or jazz. However, these categories are quite subjective  
15 and broad. A buyer cannot expect to like every offering in a particular category, even if it is a preferred category. For instance, a jazz enthusiast will not always like every new "jazz" recording just because it is categorized as "jazz". In addition, many songs may be hard to categorize. For example, one person may think a particular song is a "rock" song, while another person thinks it is a "rhythm & blues" song. The lack of  
20 consistent and repeatable classifications make searching for music using these traditional categories difficult.

Therefore, people frequently read reviews of a musical CD or other offering in order to determine whether or not they would like to purchase a particular CD. After purchasing the CD, buyers are frequently disappointed with their purchase because  
25 their subjective opinion of the music quite naturally differed from the reviewer's opinion.

This problem has not improved with the advent of digital music downloads. The Internet offers people more choices of music, but along with it, it also offers more reviewers giving subjective opinions of the music. People are still frequently  
30 disappointed with their music purchases.

Several methods have been developed to help buyers find music that they want to purchase. General entertainment Websites provide options to search for music by artist name and/or song title, and allow browsing through predefined music categories. Once the buyer finds something he wants, he can link to another site to purchase the music. However, these sites do little more than list what is available, and provide basic search capabilities.

Websites that track the popularity of downloaded music have also been developed. These sites rate and compute the most popular downloads, and provide links for potential buyers to link to sites selling the music. While these sites offer some additional information for buyers searching for downloadable music, they only provide for those who are looking for "popular" music as opposed to finding something that matches their own personal tastes and preferences.

To account for personal tastes and preferences, a Website has been developed that provides tools for "learning" a potential music buyer's tastes. However, the Website is not using objective classifications, but instead builds a "clustering" database using a technique referred to as "collaborative filtering." From the database, the Website can determine general trend information such as "People who like Artist A also like Artist B." Such analysis, however, only uncovers popular trends. As the number of songs on the Web increases, this method will be prone to confusion since the number of possible correlations becomes endless. Furthermore, the collaborative filtering technique does not allow the introduction of new or previously unheard music. It is merely a "black box" that reflects the choices of others, but not why such choices were made. In addition, the black box becomes relatively unstable with large inputs.

Search engines for MP3 files have been developed to help a user find a particular song or style of music. The search engines attempt to describe and categorize the Web's massive supply of digital downloads. Musical experts are hired to describe every new track and compare it to a well-known band. Using these search engines, users can find music that is subjectively similar to music that they know they like. However, the results are subjective. Users may or may not agree with the

experts' opinions. It is a subjective method of evaluating music, and while a definite improvement over simple keyword searching, the results can vary depending on the reviewer. Also, as the number of MP3 files online increase dramatically, this method will require additional music reviewing staff to maintain the database and provide users with current information. Consequently, the domain of existing music (such as music from certain time periods such as 1960, 1970 and 1980) may not be classified for a relatively long period of time, if ever.

In view of the foregoing, it can be appreciated that a substantial need exists for a system and method for objectively categorizing music in a consistent, repeatable manner. There is a need for a system that can manage the massive number of music downloads available to a user on the Internet.

#### SUMMARY OF THE INVENTION

One embodiment of the invention comprises a method and apparatus for categorizing music. A digital signal representing music is received. Descriptors are generated using said digital signal. The music is categorized using said descriptors.

With these and other advantages and features of the invention that will become hereinafter apparent, the nature of the invention may be more clearly understood by reference to the following detailed description of the invention, the appended claims and to the several drawings attached herein.

#### BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram for a system suitable for practicing one embodiment of the invention.

FIG. 2 is a block diagram for a computer system suitable for practicing one embodiment of the invention.

FIG. 3 is a block flow diagram of steps performed by a music classification module in accordance with one embodiment of the invention.

FIG. 4 is a block flow diagram of steps to generate descriptors in accordance with one embodiment of the invention.

FIG. 5 is a block flow diagram of steps to create mathematical descriptions in accordance with one embodiment of the invention.

Fig. 6 illustrates a statistical modeling by wavelets in accordance with one embodiment of the invention.

## 5            DETAILED DESCRIPTION

The embodiments of the invention comprise a method and apparatus to categorize music. The amount of digital music on the Internet and elsewhere is increasing. Consumer desire for such music is also increasing. There is therefore a need for an objective music classification scheme. Presently, music is classified using  
10 the names of the artists, the year it was produced and the general genre of the music, such as pop, rock or jazz. However, with the increasing amount of available and stored music, such subjective categories are not effective in grouping similarly sounding music.

People tend to like a certain type or style of music. When they search for new  
15 music, it is a certain sound they are looking for, not a genre. Therefore, there is a need to be able to classify similar-sounding music together. There is a need for an objective classification scheme that uses the music itself in determining the class, instead of the current method of using subjective criteria and/or derived psycho-acoustic properties of the song like beat, rhythm or tempo. The system and method of  
20 the present invention provides an objective classification scheme that can be used to search for new music over a network (e.g., the Internet or WWW) and organize personal collections on a PC or portable playback devices.

Digital music may be music that is stored on an electronic device. There are a number of known audio storage formats, including the popular MP3 format. MP3  
25 was developed under the sponsorship of the Moving Picture Experts Group (MPEG) as a standard technology and format for compressing a sound sequence into a very small file (about one-twelfth the size of the original file) while preserving the original level of sound quality when it is played. New audio storage methodologies under

development such as MPEG-4 and SDMI, as well another known formats are considered to be within the scope of the present invention.

MP3 files are usually download-and-play files. However, digital music also includes streaming sound, which is sound that is played as it arrives, or alternatively a  
5 sound recording (such as a WAV file) that doesn't start playing until the entire file has arrived. Support for streaming sound may require a plug-in player or come with a Web browser. Digital music as used in the present invention is intended to cover any type of digital audio, including streaming sound.

Digital music is just like any other form of data, such as astronomical image  
10 data. As the amount of scientific data has increased, researchers have developed new statistical methods for extracting important information from the data quickly and accurately. These same digital signal processing techniques can be used to extract information about digital music. The "data" that represents music is processed into intermediate data products that isolate the essential information content of the music.

15 Therefore, using the latest techniques in digital signal processing, the data can be decomposed into its most common components that can then be used to mathematically characterize the music. This mathematical description of the digital music can used to objectively compare different pieces of music. Moreover, these characteristics can be used as a method of grouping similar music, and thereby  
20 establish an objective classification scheme. Trends between different songs can be identified using the mathematical description. The system and method of the present invention can be given new songs and be able to identify other music that sounds like the new song using the mathematical description.

It is worthy to note that any reference in the specification to "one  
25 embodiment" or "an embodiment" means that a particular feature, structure or characteristic described in connection with the embodiment is included in at least one embodiment of the invention. The appearances of the phrase "in one embodiment" in various places in the specification are not necessarily all referring to the same embodiment.



Referring now in detail to the drawings wherein like parts are designated by like reference numerals throughout, there is illustrated in FIG. 1 a system suitable for practicing one embodiment of the invention. FIG. 1 is a block diagram of a communication system 100 comprising a client computer system 102 and a server computer system 106 connected via a network 104. In one embodiment of the invention, network 104 is a network capable of communicating using a variety of protocols, such as the Transport Control Protocol/Internet Protocol (TCP/IP) and File Transport Protocol (FTP) used by the Internet, and the HTTP used by the World Wide Web "WWW". Server computer system 106 is an application server, and contains one or more files containing digital data representing music. The files could be in any conventional format suitable for storing digital data for music, such as a MP3 file or a .WAV file.

FIG. 2 is a block diagram of a computer system 200 which is representative of client computer system 102 and server computer system 104, in accordance with one embodiment of the invention. Each of these blocks represents at least one such computer system. Although only one each of client computer system 102 and server computer system 104 are shown in FIG. 1, it is well known in the art that multiple computer systems can be available and still fall within the scope of the invention. Further, it is also well known in the art that a distributed architecture in which more than one computer system performs each function is entirely equivalent.

In one advantageous embodiment of the invention, computer system 200 represents a portion of a processor-based computer system. Computer system 200 includes a processor 202, an input/output (I/O) adapter 204, an operator interface 206, a memory 210 and a disk storage 218. Memory 210 stores computer program instructions and data. Processor 202 executes the program instructions, and processes the data, stored in memory 210. Disk storage 218 stores data to be transferred to and from memory 210. I/O adapter 204 communicates with other devices and transfers data in and out of the computer system over connection 224. Operator interface 206 interfaces with a system operator by accepting commands and providing status information. All these elements are interconnected by bus 208, which allows data to

be intercommunicated between the elements. I/O adapter 204 represents one or more I/O adapters or network interfaces that can connect to local or wide area networks such as, for example, the network described in FIG. 1. Therefore, connection 224 represents a network or a direct connection to other equipment.

5           Processor 202 can be any type of processor capable of providing the speed and functionality required by the embodiments of the invention. For example, processor 202 could be a processor from a family of processors made by Intel Corporation, Motorola, AMD, Compaq Corporation or others.

10           For purposes of this application, memory 210 and disk 218 are machine readable mediums and could include any medium capable of storing instructions adapted to be executed by a processor. Some examples of such media include, but are not limited to, read-only memory (ROM), random-access memory (RAM), programmable ROM, erasable programmable ROM, electronically erasable programmable ROM, dynamic RAM, magnetic disk (e.g., floppy disk and hard drive),  
15   optical disk (e.g., CD-ROM), optical fiber, electrical signals, lightwave signals, radio-frequency (RF) signals and any other device or signal that can store digital information. In one embodiment, the instructions are stored on the medium in a compressed and/or encrypted format. As used herein, the phrase “adapted to be executed by a processor” is meant to encompass instructions stored in a compressed  
20   and/or encrypted format, as well as instructions that have to be compiled, interpreted or installed by an installer before being executed by the processor. Further, system 200 may contain various combinations of machine readable storage devices through other I/O controllers, which are accessible by processor 202 and which are capable of storing a combination of computer program instructions and data.

25           I/O adapter 204 includes a network interface that may be any suitable means for controlling communication signals between network devices using a desired set of communications protocols, services and operating procedures. As mentioned previously, in one embodiment of the invention, I/O adapter 204 utilizes the transport control protocol (TCP) of layer 4 and the internet protocol (IP) of layer 3 (often  
30   referred to as “TCP/IP”). I/O adapter 204 also includes connectors for connecting I/O

adapter 204 with a suitable communications medium (e.g., connection 224). Those skilled in the art will understand that I/O adapter 204 may receive communication signals over any suitable medium such as twisted-pair wire, co-axial cable, fiber optics, radio-frequencies, and so forth.

5           Memory 210 is accessible by processor 202 over bus 208 and includes an operating system 216, a program partition 212 and a data partition 214. Program partition 212 may be a single or multiple program partition which stores and allows execution by processor 202 of program instructions that implement the functions of each respective system described herein. Data partition 214 is accessible by processor  
10   202 and stores data used during the execution of program instructions.

          In one embodiment of the invention, program partition 212 contains program instructions that are used to categorize music by analyzing a digital signal containing information representing the music. These program instructions will be referred to herein collectively as a “music categorization module.” The music categorization  
15   module utilizes digital signal processing to create a mathematical description of the music. The mathematical description is used to classify music based on the actual music itself versus subjective perceptions of the music. The operation of systems 100, 200 and a music categorization module will be described with reference to FIGS. 3-6.

20           FIG. 3 is a block flow diagram of steps performed by a music classification module in accordance with one embodiment of the invention. As shown in FIG. 3, a digital signal representing music is received at step 302. Descriptors are generated using the digital signal at step 304. The music is categorized using the descriptors at step 306.

25           The received digital signal representing music can be in any number of conventional formats. For example, a song can be converted from an analog format to a digital format, such as the raw .WAV format, the MP3 format and the SDMI format. These formats represent audio file types that have been accepted as a viable interchange medium between different computer platforms, allowing content

developers to freely move audio files between platforms for various purposes, such as processing.

FIG. 4 is a block flow diagram of steps to generate descriptors in accordance with one embodiment of the invention. The term “descriptors” are used herein to  
5 identify information used to categorize music, such as data, coefficients, values, parameters, mathematical descriptions, and so forth. As shown in FIG. 4, mathematical descriptions of the digital signal are created at step 402. The mathematical descriptions are represented as vectors at step 404. The vectors are clustered into statistically significant groups at step 406.

10 FIG. 5 is a block flow diagram of steps to create mathematical descriptions in accordance with one embodiment of the invention. In this embodiment of the invention, wavelets are used as the basis for the mathematical description. As shown in FIG. 5, a spectrogram is formed from the digital signal at step 502. The spectrogram is renormalized in frequency space at step 504. A wavelet image is  
15 generated using a dual transform analysis of the spectrogram at step 506. The coefficients are selected from the wavelet image at step 508.

A spectrogram is a data file containing the power spectrum of the Fast Fourier Transform as a function of time. In one embodiment of the invention, the spectrogram is formed by taking  $\Delta t$  segments of the song ( $\Delta t$  is user definable) and  
20 computing the Fast Fourier Transform. The square of the amplitude, which is the power spectrum, is kept.

The digital signal representing an input waveform can be decomposed into various components using a number of methods, such as a Fast Fourier Transform (cosines & sines), a wavelet transform (wavelet packets), Cosine packets, any  
25 orthonormal based transform methods or any principal component analysis transform methods. With respect to wavelet transforms, a number of different wavelet packets can be generated, such as Daubechies, Symmlet, Coiflet or “Mexican Hat” wavelet packets.

In one embodiment of the present invention, wavelets are used as the basis for  
30 the mathematical description. It can be appreciated, however, that other descriptors

can be used and still fall within the scope of the invention. For example, any of the methods or techniques described above can be used as a basis for the mathematical description, and still fall within the scope of the invention.

A wavelet is a mathematical function useful in many different digital signal processing applications. For example, wavelets are used in image compression applications by analyzing an image and converting it into a set of mathematical expressions that can then be decoded by the receiver. Wavelet functions cut up data into different frequency components, and then study each component with a resolution matched to its scale. Wavelets are specifically designed to decompose data into their main, orthogonal components.

More particularly, a wavelet is an orthonormal basis that is localized in both space and frequency. The "mother wavelet" has compactness in space and frequency and should integrate to zero. An input signal is decomposed into an orthonormal set of scaled wavelets via translation and dilation. The size or coefficient of these scaled wavelets is stored and the highest values provide an exponential compression of the information in the signal, as illustrated in FIG. 6.

Fig. 6 illustrates a statistical modeling by wavelets in accordance with one embodiment of the invention. As shown in Fig. 6, the doppler function 610, is decomposed into a series of numbers at different resolutions. These are the coefficients d1 through d10. Only the highest fraction of these coefficients need to be saved in order to accurately reproduce the original function. The coefficients can then be used to classify it, and search for other functions with similar coefficients.

One embodiment of the invention decomposes a relatively complicated input signal into a set of coefficients in different levels (e.g., as shown in FIG. 6). Each level represents a factor of 2 dilation in the mother wavelet (i.e., twice as big at each level down). At each point in each level, the size or coefficient of the wavelet is generated as needed to match the input signal at that particular point or position. If the process were to be reversed (i.e., only keep the largest N coefficients, and place the wavelet, scaled appropriately, at the position of each of these large coefficients), it can be appreciated that an acceptable reproduction of the original input image in both

frequency and space can be recovered. The N coefficients are a condensed representation of the data.

Referring again to FIG. 5, a spectrogram is formed from the digital signal at step 502. This can be accomplished by taking intervals of time sections and  
5 performing a Fast Fourier Transform of these sections. The components may be limited to real components, or may include imaginary or phase information as well. The spectrogram is renormalized in frequency space at step 504. The spectrogram is split in frequency space, and a dual wavelet transform analysis is performed at step 506 to form a wavelet image. The term "dual wavelet transform analysis" refers to  
10 performing a wavelet transform analysis on each part (e.g., above and below the frequency split). By splitting the spectrogram, the emphasis on harmonics is enhanced, which often occurs at higher frequencies and determines the instrumentation used in the music. This may be performed by, for example, using Coiflet, Symmlets, Daubechies (e.g., Daubechies 2, Daubechies 4 and Daubechies 8),  
15 Cosine or Mexican Hat packets. A particular method may be selected based on the desired smoothness of the resulting wavlets. For example, each mother wavelet (e.g., Coiflet, Symmlets, Daubechies) has an associated smoothness.

Although a dual wavelet transform analysis is shown in this embodiment of the invention, it can be appreciated that other wavelet transform analysis may be  
20 applied and still fall within the scope of the invention. For example, a wavelet transform may be performed on all segments (e.g., more than two images) if desired.

The coefficients are selected from the wavelet image at step 508. In one embodiment of the invention, the top N coefficients from the wavelet image are selected. For example, N may be equal to 1000 which would represent approximately  
25 0.1% of the input data. The selection criteria may vary for each application, and may include such criteria as selecting the N highest magnitude, N with highest standard deviation or N with highest magnitude and standard deviation.

The coefficients, or other musical descriptors, are calculated and saved for various digital music. By way of contrast, conventional classification techniques use  
30 humans to classify music by ear, or they use psycho-acoustic parameters like beat,

rhythm or tempo. The latter items are computed from the music, but typically only use 3 numbers. Once a large database of coefficients or musical descriptions is created in accordance with the embodiments of the invention, the music may be classified.

5           In one embodiment, existing categories of music are used. These existing categories are typically known genres, such as rock or jazz. In this embodiment, coefficients for each category are determined, and music that has similar coefficients is classified as being in that category. For example, analysis of music that has previously been classified as "rock" may reveal that rock music only has large d8 and  
10       d10 coefficients. By making this determination, new music that has large d8 and d10 coefficients can be classified as rock. Once a scheme is established, any new music that is analyzed by the method and system of the present invention can immediately be related to other music via these standard coefficients.

          The determination of coefficients for an existing category may be made in  
15       several ways. A neural network may be created with R middle layers defining common properties of song musical descriptions in each of the existing categories. Alternatively, a Bayes Network may be used to define common properties of songs in each of the existing categories. Other methods are known to those skilled in the art, and are intended to be within the scope of the present invention.

20           In another embodiment of the present invention, natural groupings or clusters are determined instead of using pre-existing categories. In this embodiment, music is categorized as belonging to a class with similar coefficients. Instead of forcing the music into a pre-existing category, categories are created based on the music itself. By creating new grouping using analysis of the music itself, the classification scheme  
25       is even more precise. For example, in one embodiment of the invention Bayes Networks are used to determine the natural clustering of the coefficients to define new genres that are more natural for the music itself.

          For this embodiment, there are no pre-determined classifications. Instead, the analysis creates groups that are used to identify music that sounds similar. One  
30       method of creating the groups is to represent each song as a vector in the N-

dimensional Fourier/wavelet space. Known mathematical algorithms are used to cluster the vectors into statistically significant groups with no pre-determined size, shape or orientation in the N-dimensional space. These new groupings of song vectors are the basis for a new objective classification scheme. In this embodiment, the music is allowed to cluster itself in N-dimensional space.

Many methods can be used to group the songs. For example, k-means, mixture modeling, adaptive and non-adaptive kernel density estimation, voronoi tessellation, or matched filtering may be used. Other methods are known to those skilled in the art, and are intended to be within the scope of the present invention. These groupings of song vectors can then be used in Neural Network and Bayes Network instead of the pre-defined classes, as discussed above.

For example, one embodiment of the invention utilizes mixture modeling analysis to group the songs. A mixture model is the use of k-kernels which are fit to the data. This is a non-parametric analysis and typically a gaussian kernel is used. More particularly, k-gaussians (which are allowed to each change shape, position and size) are fit the point data in N dimensions. These gaussians adaptively smooth the data providing a probability density map of this N dimensional space, which can then be searched, or thresholded, for peaks. These peaks become the new classes, or rather the size and shape of these peaks assist in formulating new classes.

In yet another embodiment of the invention to categorize or group music, each individual person may be considered a separate category or bin. In essence, each person represents a personal classification based on songs or music identified by, or associated with, the individual. Songs could then be classified or grouped according to each person, and new songs can be pushed to various people based on a set of descriptors associated or formulated for each person.

Once categories are established, new songs can be added to a database. The musical description, or coefficients, of the new song are compared to the regions that the Neural Network and/or Bayes Network defined for the pre-existing classes, natural groupings or personal groupings. The song is then assigned a mathematical likelihood of being a member of each of these classes or groupings. The highest



likelihood is assigned the class or grouping of the song, thus objectively classifying a new song. Songs can have high likelihoods of being in multiple classes or groupings.

In an alternative embodiment, supplemental information can be added to the classification process. By storing supplemental information with the music data, a  
5 profile of the listener can be generated and provided to advertisers. Examples of supplemental information include beat, rhythm, existing genres, other songs people like, demographic information (e.g., age, income, gender, location, etc.), and so forth. The combination of coefficients and supplemental information can then be clustered in the  $N$  (coefficients) +  $M$  (supplemental) dimensional space. The algorithms  
10 discussed previously, such as k-mean, can be used for the classification process. The distance metric, i.e., the desired distance between two vectors in this  $N + M$  dimensional space, would be defined according to a particular application.

Once a music classification scheme is established, whether using pre-existing classes, new natural clusters or personal groupings, many search options become  
15 possible. For example, a user can search for all music that sounds like a particular group or class of songs, or even all music that sounds most dissimilar to a particular song. It is possible to do very specific searches, such as all music by The Beatles that sound like "Hey Jude."

As one can imagine, there are many uses for the system and method of the  
20 present invention. One use is to generate a playlist based on the objective classifications of digital music. A personal playlist can be generated based on classifications and downloaded from a network, such as the Internet. A fully automated, personalized Streaming Radio Playlist can be generated. Music on electronic devices that store and play digital music can be managed using the  
25 objective classification scheme of the present invention.

One advantageous use of the system and method of the present invention is to search for music. The Internet or other network can be searched for new music that sounds similar to a particular person, song, group of songs, genre (existing or natural) or songs of a particular known band. The system and method of the present invention  
30 can also be used offline to search inventory in records stores to find new music that

sounds similar to a song. For this use, a record store may use a kiosk for the searching system.

A recording studio may use the system and method of the present invention to help identify the next big hit based on an objective analysis of past hits. Recording studios may also use the system and method of the present invention to automate the selection of a similar song to attach free to the end of a CD as a sales tool.

Musicians may use the system and method of the present invention to generate new music that will be more likely to reach a particular audience based on objective classification of the music itself.

The system and method of the present invention may be used to provide purchasing information based on sales. New music may be offered for sale to record stores and the available selection will be based on the objective classification of new music and it's match to the "sales profile" of that particular retailer. This may be used by both online and physical stores. The system and method of the present invention may also be used to suggest new music to a customer based on current and/or past purchases.

The system and method of the present invention may also be used by a "webcrawler" or "bot" to establish a profile based on a person's musical library and constantly search the Web for new music that matches the profile. The bot may offer samples to the user, and provide methods the user to download or purchase any found music.

Although various embodiments are specifically illustrated and described herein, it will be appreciated that modifications and variations of the present invention are covered by the above teachings and within the purview of the appended claims without departing from the spirit and intended scope of the invention. For example, although the embodiments of the invention implement the functionality of the processes described herein in software, it can be appreciated that the functionality of these processes may be implemented in hardware, software, or a combination of hardware and software, using well-known signal processing techniques. In another example, the embodiments were described using a communication network. A

communication network, however, can utilize an infinite number of network devices configured in an infinite number of ways. The communication network described herein is merely used by way of example, and is not meant to limit the scope of the invention.

**CLAIMS:**

1. A method of categorizing music, comprising:  
receiving a digital signal representing music;  
generating descriptors using said digital signal; and  
categorizing said music using said descriptors.
2. The method of claim 1, wherein said generating comprises:  
creating mathematical descriptions of said digital signal;  
representing said mathematical descriptions as vectors; and  
clustering said vectors into statistically significant groups.
3. The method of claim 2, wherein said mathematical descriptions comprise  
wavelet coefficients.
4. The method of claim 3, wherein said wavelet coefficients are created using at  
least one technique of a group comprising Coiflet, Symmlets, Daubechies, Cosine  
packets or Mexican Hat.
5. The method of claim 2, wherein said creating comprises:  
forming a spectrogram for said digital signal;  
renormalizing said spectrogram in frequency space;  
generating a wavelet image using a wavelet transform analysis of said  
spectrogram; and  
selecting coefficients from said wavelet image.
6. The method of claim 5, wherein said wavelet transform analysis is a dual  
wavelet transform analysis.

7. The method of claim 1, wherein said categorizing comprises:  
generating a set of descriptors for each of a plurality of predetermined  
categories;  
comparing said descriptors to each set of descriptors; and  
5 assigning said music to at least one of said predetermined categories in  
accordance with said comparison.
8. The method of claim 7, wherein said comparing said descriptors is performed  
using a technique from a group comprising Neural network and Bayes network.
9. The method of claim 1, wherein said categorizing comprises:  
generating a previous set of descriptors to form a category;  
comparing said descriptors to said set of descriptors; and  
assigning said music to said category in accordance with said comparison.
10. The method of claim 9, wherein said comparing said descriptors is performed  
using a technique from a group comprising Neural network and Bayes network.
11. The method of claim 9, wherein said previous set of descriptors is generated  
using music associated with a particular person.
12. The method of claim 11, wherein said comparing said descriptors is performed  
using a technique from a group comprising Neural network and Bayes network.
13. The method of claim 2, wherein said clustering comprises:  
receiving supplemental information for said music; and  
clustering said vectors and said supplemental information into statistically  
significant groups.

14. The method of claim 13, wherein said vectors and said supplemental information are clustered in  $N + M$  dimensions, utilizing at least one technique from a group comprising k-means, mixture modeling, adaptive kernel density estimation, non-adaptive kernel density estimation, voronoi tessellation and matched filtering.
15. The method of claim 2, wherein said vectors are clustered utilizing at least one technique from a group comprising k-means, mixture modeling, adaptive kernel density estimation, non-adaptive kernel density estimation, voronoi tessellation and matched filtering.
16. A method of categorizing music, comprising:  
receiving a digital signal representing music from a first file having a first size;  
compressing said digital signal using a set of descriptors to form a second file  
5 having a second size smaller than said first size; and  
categorizing said music using said descriptors.
17. A machine-readable medium whose contents cause a computer system to categorize music, comprising:  
receiving a digital signal representing music;  
generating descriptors using said digital signal; and  
5 categorizing said music using said descriptors.
18. The machine-readable medium of claim 17, wherein said generating comprises:  
creating mathematical descriptions of said digital signal;  
representing said mathematical descriptions as vectors; and  
5 clustering said vectors into statistically significant groups.

19. The machine-readable medium of claim 18, wherein said mathematical descriptions comprise wavelet coefficients.
20. The machine-readable medium of claim 19, wherein said wavelet coefficients are created using at least one technique of a group comprising Coiflet, Symmlets, Daubechies, Cosine packets or Mexican Hat.
21. The machine-readable medium of claim 18, wherein said creating comprises:  
forming a spectrogram for said digital signal;  
renormalizing said spectrogram in frequency space;  
generating a wavelet image using a wavelet transform analysis of said  
5 spectrogram; and  
selecting coefficients from said wavelet image.
22. The machine-readable medium of claim 21, wherein said wavelet transform analysis is a dual wavelet transform analysis.
23. The machine-readable medium of claim 17, wherein said categorizing comprises:  
generating a set of descriptors for each of a plurality of predetermined  
categories;  
5 comparing said descriptors to each set of descriptors; and  
assigning said music to at least one of said predetermined categories in  
accordance with said comparison.
24. The machine-readable medium of claim 23, wherein said comparing said descriptors is performed using a technique from a group comprising Neural network and Bayes network.

25. The machine-readable medium of claim 17, wherein said categorizing comprises:

- generating a previous set of descriptors to form a category;
- comparing said descriptors to said set of descriptors; and
- 5 assigning said music to said category in accordance with said comparison.

26. The machine-readable medium of claim 25, wherein said comparing said descriptors is performed using a technique from a group comprising Neural network and Bayes network.

27. The machine-readable medium of claim 25, wherein said previous set of descriptors is generated using music associated with a particular person.

28. The machine-readable medium of claim 27, wherein said comparing said descriptors is performed using a technique from a group comprising Neural network and Bayes network.

29. The machine-readable medium of claim 18, wherein said clustering comprises:  
receiving supplemental information for said music; and  
clustering said vectors and said supplemental information into statistically significant groups.

30. The machine-readable medium of claim 29, wherein said vectors and said supplemental information are clustered in  $N + M$  dimensions, utilizing at least one technique from a group comprising k-means, mixture modeling, adaptive kernel density estimation, non-adaptive kernel density estimation, voronoi tessellation and  
5 matched filtering.

31. The machine-readable medium of claim 18, wherein said vectors are clustered utilizing at least one technique from a group comprising k-means, mixture modeling,



adaptive kernel density estimation, non-adaptive kernel density estimation, voronoi tessellation and matched filtering.

32. A machine-readable medium of categorizing music, comprising:  
receiving a digital signal representing music from a first file having a first size;  
compressing said digital signal using a set of descriptors to form a second file  
5 having a second size smaller than said first size; and  
categorizing said music using said descriptors.
33. A method to search for music, comprising:  
receiving a request for a first set of music based on a second set of music, said  
second set of music having been identified by a second set of descriptors using  
wavelet analysis;  
5 identifying a first set of descriptors for said first set of music using wavelet  
analysis;  
comparing said first set of descriptors with said second set of descriptors; and  
retrieving said first set of music in accordance with said comparison.
34. An apparatus to categorize music, comprising:  
means for receiving a digital signal representing music;  
means for generating descriptors using said digital signal; and  
means for categorizing said music using said descriptors.
35. A system to categorize music, comprising:  
a network;  
a computer system connected to said network to receive music in a digital  
format, and to identify a first set of descriptors for said music using wavelet  
5 analysis;

a memory to store said first set of descriptors; and  
a search module to search for said first set of descriptors in said memory.

36. The system of claim 35, wherein said search module searches for said first set of descriptors using a second set of descriptors.

37. The system of claim 35, further comprising a music categorization module to categorize said set of descriptors in accordance with at least one of a group comprising predetermined categories, natural groupings and personal groupings.

38. The system of claim 35, further comprising a music categorization module that categorizes said set of descriptors in accordance with at least one of a group comprising Neural network, Bayes network, k-means, mixture modeling, adaptive kernel density estimation, non-adaptive kernel density estimation, voronoi tessellation  
5 and matched filtering.

39. The method of claim 33, wherein said descriptors are objective descriptors.

100

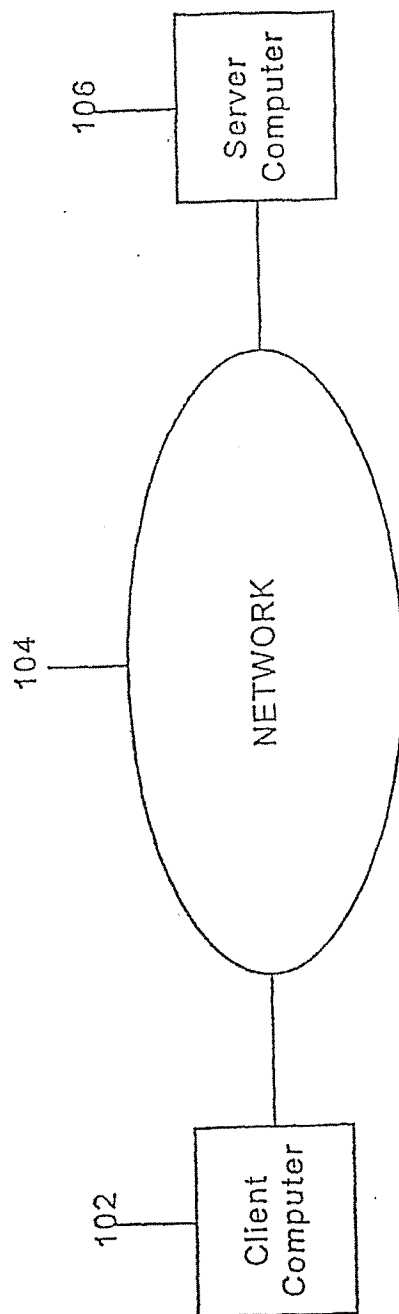


FIG. 1

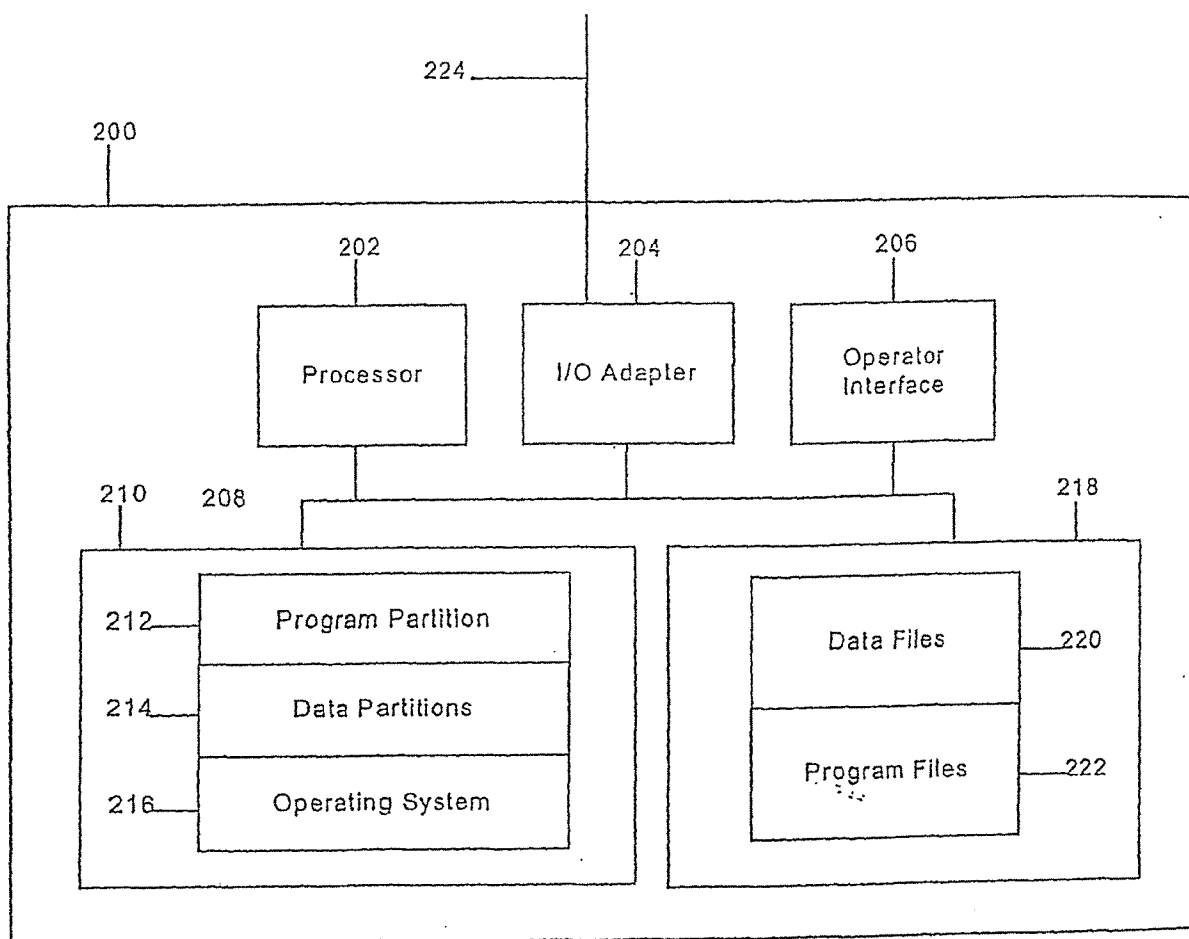


FIG. 2

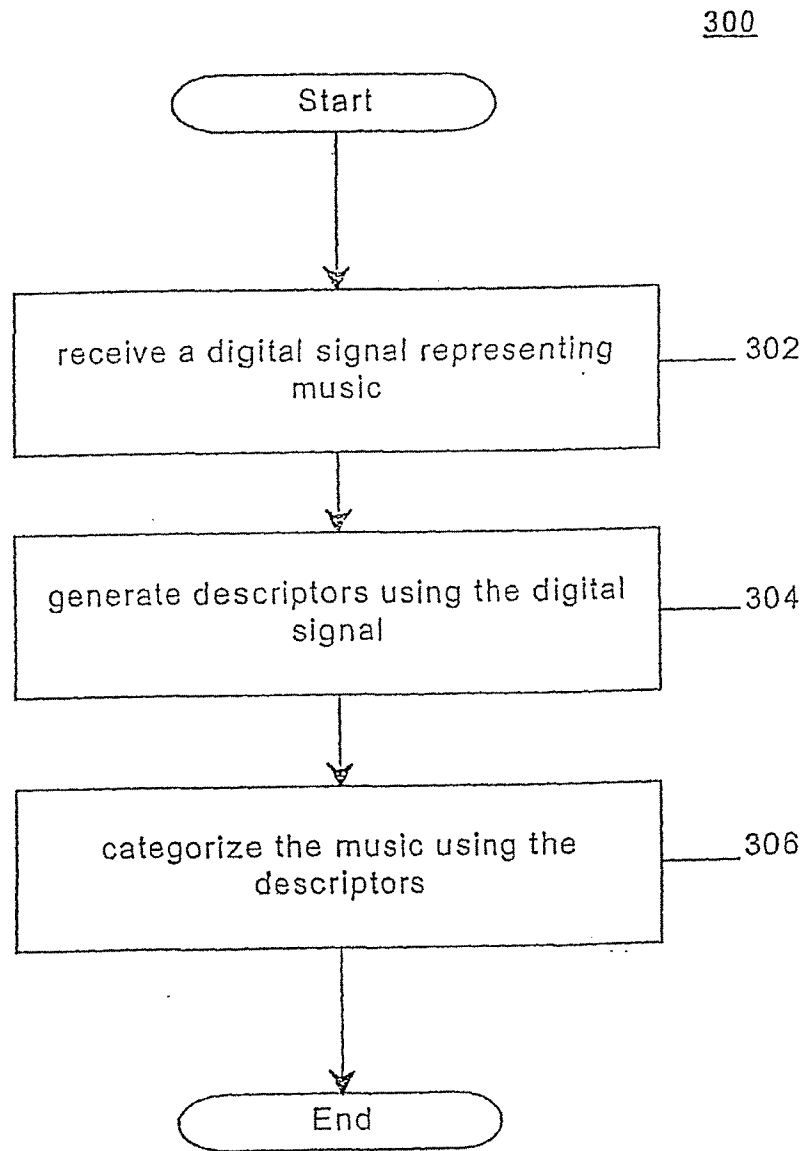


FIG. 3

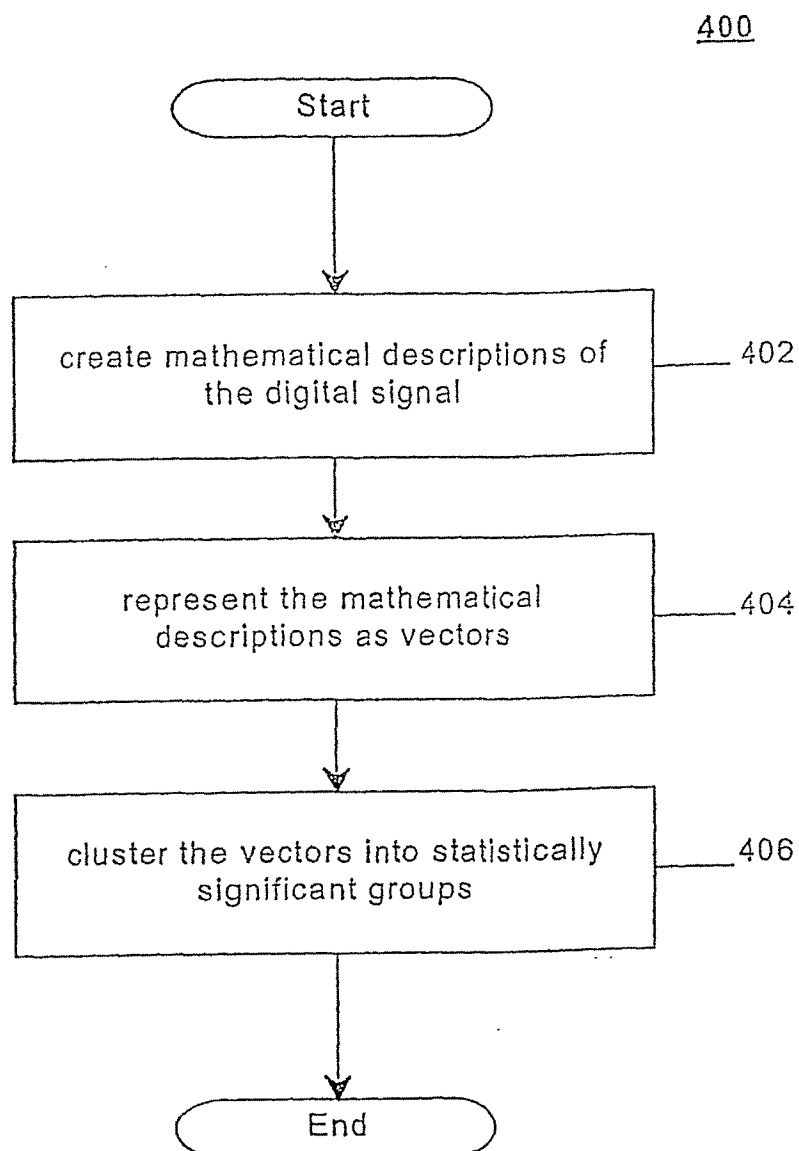


FIG. 4

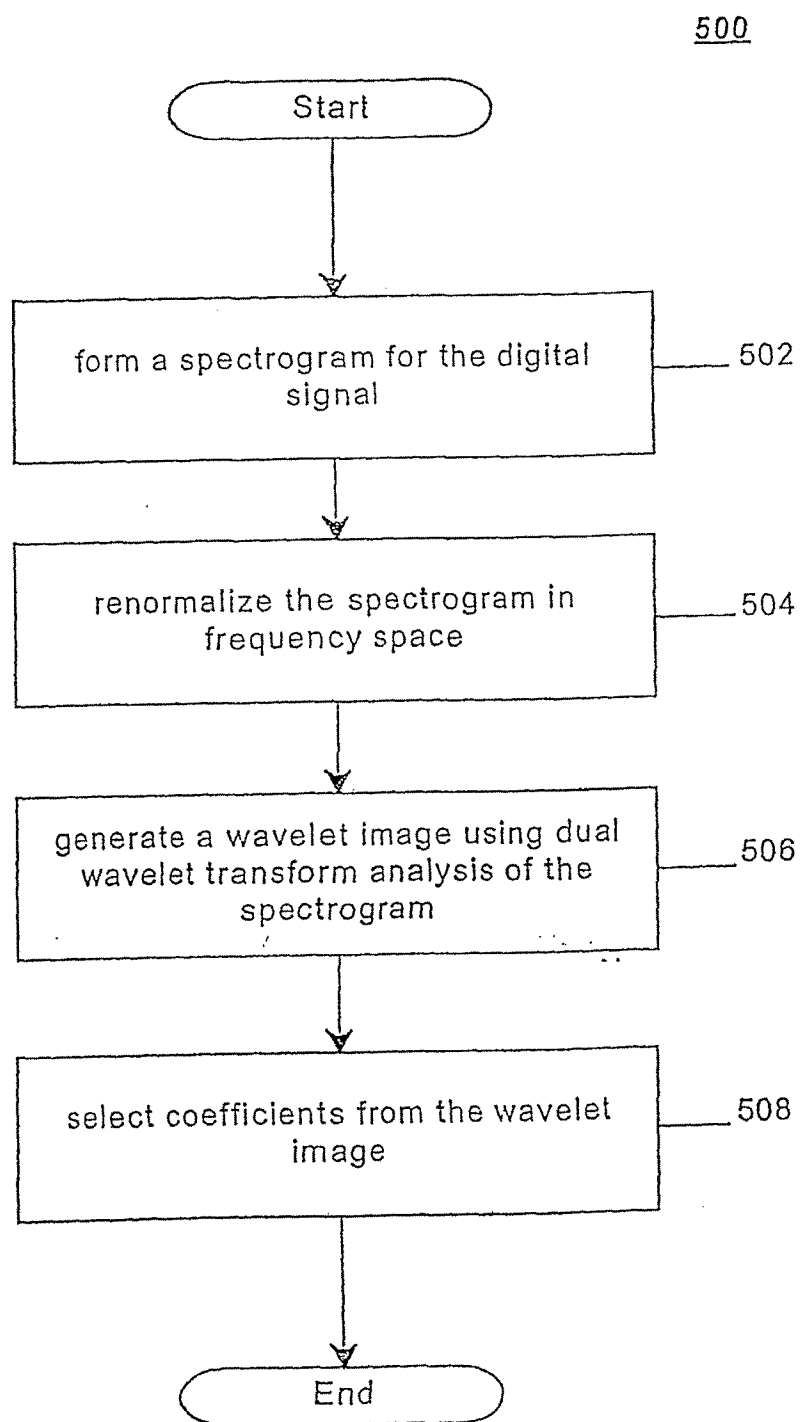


FIG. 5

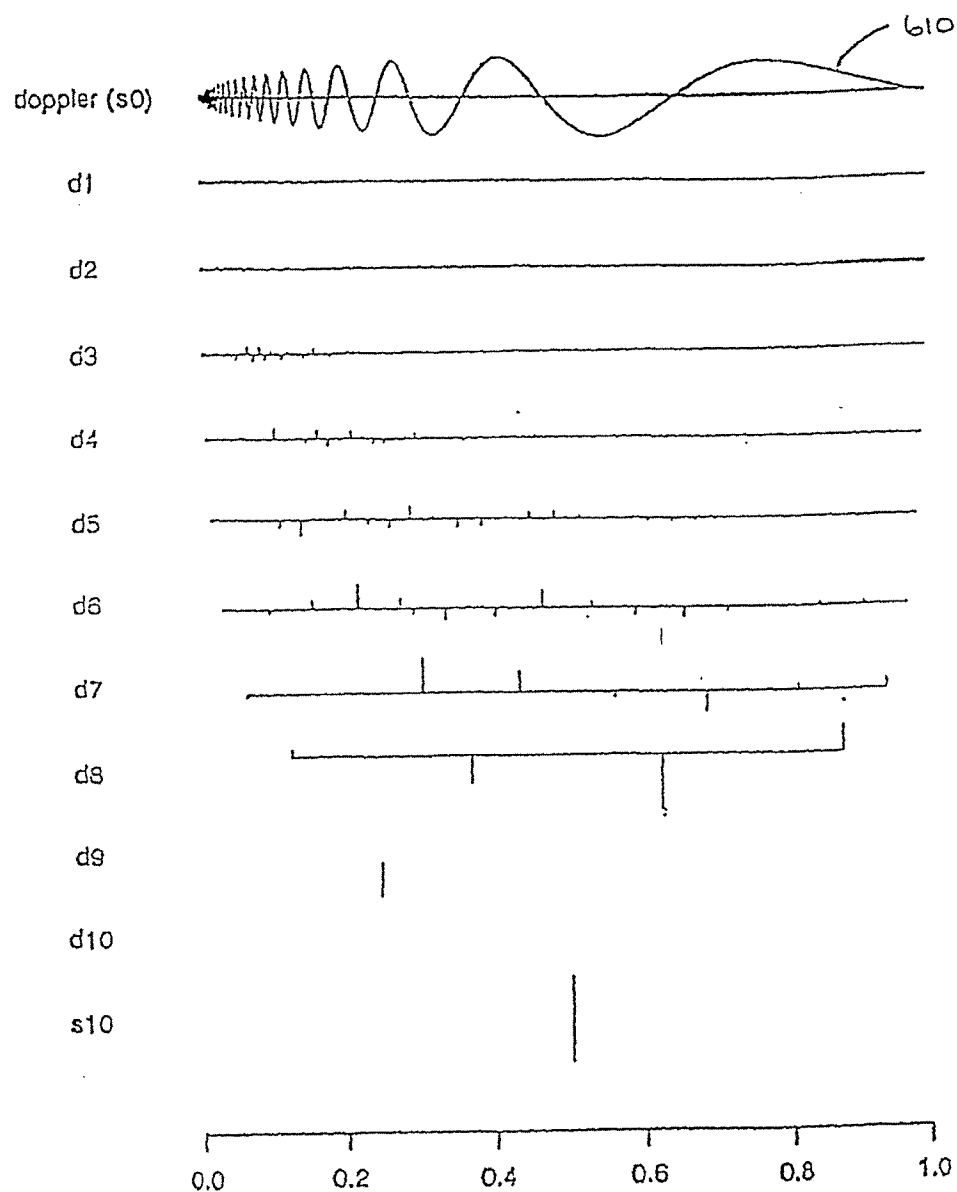


FIG. 6